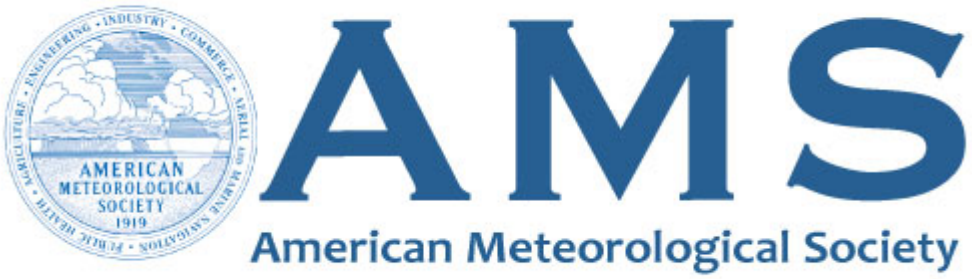

Supplemental Material

Journal of Hydrometeorology

Decisive Atmospheric Circulation Indices for July–August Precipitation in North China Based on Tree Models

https://doi.org/10.1175/JHM-D-19-0045.1

Supplementary

Decisive atmospheric circulation indices for July–August precipitation in North China based on tree models

Tong Xuan[1,2], Yan Zhongwei[1,2], Xia Jiangjiang[1], Lou Xiao[1]

1. Institute of Atmospheric Physics, Chinese Academy of Sciences

Beijing 100029 China

2. University of Chinese Academy of Sciences, Beijing 100049 China

1. Large-scale circulation indices

The monthly series of 88 atmospheric circulation indices (for the period 1951–2017) operationally applied at the China Meteorological Administration are listed as follows:

1) Northern Hemisphere Subtropical High Area Index

2) North African Subtropical High Area Index

3) North African-North Atlantic-North American Subtropical High Area Index

4) Indian Subtropical High Area Index

5) Western Pacific Subtropical High Area Index

6) Eastern Pacific Subtropical High Area Index

7) North American Subtropical High Area Index

8) Atlantic Subtropical High Area Index

9) South China Sea Subtropical High Area Index

10) North American-Atlantic Subtropical High Area Index

11) Pacific Subtropical High Area Index

12) Northern Hemisphere Subtropical High Intensity Index

13) North African Subtropical High Intensity Index

14) North African-North Atlantic-North American Subtropical High Intensity Index

15) Indian Subtropical High Intensity Index

16) Western Pacific Subtropical High Intensity Index

17) Eastern Pacific Subtropical High Intensity Index

18) North American Subtropical High Intensity Index

19) North Atlantic Subtropical High Intensity Index

20) South China Sea Subtropical High Intensity Index

21) North American-North Atlantic Subtropical High Intensity Index

22) Pacific Subtropical High Intensity Index

23) Northern Hemisphere Subtropical High Ridge Position Index

24) North African Subtropical High Ridge Position Index

25) North African-North Atlantic-North American Subtropical High Ridge Position Index

26) Indian Subtropical High Ridge Position Index

27) Western Pacific Subtropical High Ridge Position Index

28) Eastern Pacific Subtropical High Ridge Position Index

29) North American Subtropical High Ridge Position Index

30) Atlantic Sub Tropical High Ridge Position Index

31) South China Sea Subtropical High Ridge Position Index

32) North American-North Atlantic Subtropical High Ridge Position Index

33) Pacific Subtropical High Ridge Position Index

34) Northern Hemisphere Subtropical High Northern Boundary Position Index

35) North African Subtropical High Northern Boundary Position Index

36) North African-North Atlantic-North American Subtropical High Northern

Boundary Position Index

37) Indian Subtropical High Northern Boundary Position Index

38) Western Pacific Subtropical High Northern Boundary Position Index

39) Eastern Pacific Subtropical High Northern Boundary Position Index

40) North American Subtropical High Northern Boundary Position Index

41) Atlantic Subtropical High Northern Boundary Position Index

42) South China Sea Subtropical High Northern Boundary Position Index

43) North American-Atlantic Subtropical High Northern Boundary Position Index

44) Pacific Subtropical High Northern Boundary Position Index

45) Western Pacific Sub Tropical High Western Ridge Point Index

46) Asia Polar Vortex Area Index

47) Pacific Polar Vortex Area Index

48) North American Polar Vortex Area Index

49) Atlantic-European Polar Vortex Area Index

50) Northern Hemisphere Polar Vortex Area Index

51) Asia Polar Vortex Intensity Index

52) Pacific Polar Vortex Intensity Index

53) North American Polar Vortex Intensity Index

54) Atlantic-European Polar Vortex Intensity Index

55) Northern Hemisphere Polar Vortex Intensity Index

56) Northern Hemisphere Polar Vortex Central Longitude Index

57) Northern Hemisphere Polar Vortex Central Latitude Index

58) Northern Hemisphere Polar Vortex Central Intensity Index

59) Eurasian Zonal Circulation Index

60) Eurasian Meridional Circulation Index

61) Asian Zonal Circulation Index

62) Asian Meridional Circulation Index

63) East Asian Trough Position Index

64) East Asian Trough Intensity Index

65) Tibet Plateau Region 1 Index

66) Tibet Plateau Region-2 Index

67) India-Burma Trough Intensity Index

68) Arctic Oscillation

69) Antarctic Oscillation

70) North Atlantic Oscillation

71) Pacific/North American Pattern

72) East Atlantic Pattern

73) West Pacific Pattern

74) North Pacific Pattern

75) East Atlantic-West Russia Pattern

76) Tropical-Northern Hemisphere Pattern

77) Polar-Eurasia Pattern

78) Scandinavia Pattern

79) Pacific Transition Pattern

80) 30 hPa zonal wind Index

81) 50 hPa zonal wind Index

82) Mid-Eastern Pacific 200 mb Zonal Wind Index

83) West Pacific 850 mb Trade Wind Index

84) Central Pacific 850 mb Trade Wind Index

85) East Pacific 850 mb Trade Wind Index

86) Atlantic-European Circulation W Pattern Index

87) Atlantic-European Circulation C Pattern Index

88) Atlantic-European Circulation E Pattern Index

Due to numerous missing values, we excluded 10 of the 88 indices: numbers 4, 9, 15, 20, 37, 42, 76, 79, 87 and 88 in the list above.

2. Detailed information about parameter selection

In this study, six models were constructed as follows: (1) Control Trial 1 (CT1), constructed by the common random forest (RF) procedure, designed to highlight the advantage of RRF; (2) Control Trial 2 (CT2), constructed by the logistic regression, designed to show the advantage of RRF beyond the linear method; (3–5) Intermediate Trials 1–3 (IT1–3) of RRF; and finally (6) the optimal RF model (ORF). The intermediate RF models (IT1–3) were constructed recursively (e.g., the latter model (IT3) uses the results of the former models (IT1-2)). Detailed information about the parameters chosen for constructing the model is shown in Table S2.

It should be noted that these parameters are based on the ski-learn toolkit. Here, some brief introductions for these parameters are provided. The parameters “Max feature”, “Max depth”, “Min samples split”, “Max leaf nodes” and “Max leaf nodes” are unique for tree models (referred to as RF in this study). “Estimator numbers” means the number of Decision Trees used for the ensemble outcome of RF. “Random state” represents the random seeds for fixing the random states of cross validation. “Feature numbers” was used to determine the quantity of indices used to constructed model. “Penalty”, “Solver” and “Max iter” are unique for logistic regression. More detailed information about these parameters is available from https://scikit-learn.org/stable/index.html.

3. The seasonal cycles of chosen 19 stations

The seasonal cycles of precipitation of the 19 stations are shown in Figure S1. It is obvious that these 19 stations have very similar seasonal cycles with the maximum precipitation peak during July-August.

4. Long-term climate change might have influenced summer rainfall in North China

The decisive large-scale circulation factors not only have strong inter-annual variability but also exhibit long-term trends, which potentially reflect climate change. Take the North African Subtropical High Ridge Position, a decisive circulation index derived from the present study, as an example. As Figure S2 shows, the North African Subtropical High Ridge Position in July-August had a long-term southerly-moving trend, which was in-phase with the decreasing trend of summer rainfall in North China since the late 1970s.

Table S1. Atmospheric circulation indices in the three classes according to inter-correlation. Class 1 includes the indices of a maximum correlation coefficient (>0.87) with others. Class 2 includes the indices of medium correlation coefficients (0.67–0.87) among themselves. Class 3 includes the indices with a maximum correlation coefficient with others of <0.67.

Class name	Serial number
Class 1	1; 2; 3; 5 6; 7; 8; 10; 11; 12; 13; 14; 16; 17; 18; 19; 21; 22; 25; 32; 36; 43; 65; 66; 72;
Class 2	23; 24; 28; 30; 33; 34; 35; 39; 40; 41; 44; 45; 46; 49; 50; 52; 53; 55; 59; 60; 61; 62; 67; 73; 74; 75; 77; 78; 80; 81; 82; 84; 85;
Class 3	26; 27; 29; 31; 38; 47; 48; 51; 54; 56; 57; 58; 63; 64; 68; 69; 70; 71; 83; 86;

Table S2. The parameter intervals and the corresponding optimal value of parameters taken into account for the six different experiments. Values inside brackets indicate parameter intervals, underline indicates the optimal parameter value, and – means that the parameter does not need to be considered.

	Controlled trial 1	Intermediate trail 1	Intermediate trail 2	Intermediate trail 3	The optimal RF	Controlled trial 2
Max feature	[1-78] 34	[1-78] 22	[1-78] 4	[1-78] 4	[1-26] 1	–
Max depth	[1-22] 10	[1-22] 3	[1-22] 9	[1-22] 6	[1-22] 5	–
Min samples split	[2-20] 5	[2-20] 2	[2-20] 2	[2-20] 3	[2-20] 5	–
Max leaf nodes	[15-36] 18	[15-36] 3	[15-36] 20	[15-36] 15	[15-36] 14	–
Min samples leaf	[1-16] 1	[1-16] 3	[1-16] 1	[1-16] 1	[1-16] 2	–
Estimator numbers	[20-1500] 55	[20-1500] 36	[20-1500] 46	[20-1500] 40	[20-1500] 1230	[20-1500] 3

Table S2. continued

Random state	[1-100] 58	[1-100] 3	[1-100] 28	[1-100] 4	[1-100] 9	[1-100] 1
Feature numbers	–	–	–	–	[1-26] 9	[1-100] 3
Penalty	–	–	–	–	–	["newton-cg","lbfgs","liblinear","sag"] ’Lbfgs’
Solver	–	–	–	–	–	[’l1’,’l2’] ’l2’
Max iter	–	–	–	–	–	[100-500] 3

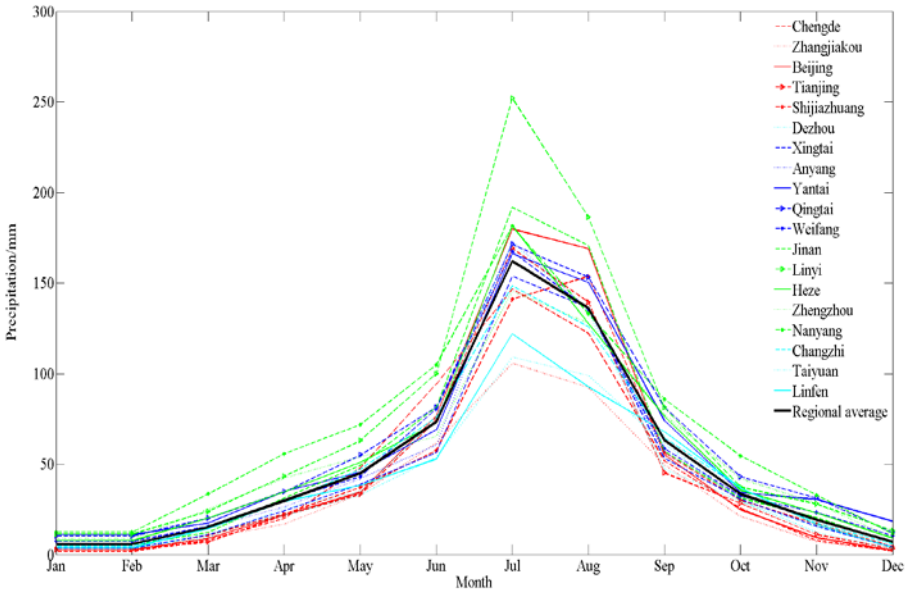

Fig.S1. Each seasonal cycle of 19 stations chosen to represent the rainfall regime in

North China and the average seasonal cycle of the total 19 stations (Black solid line)

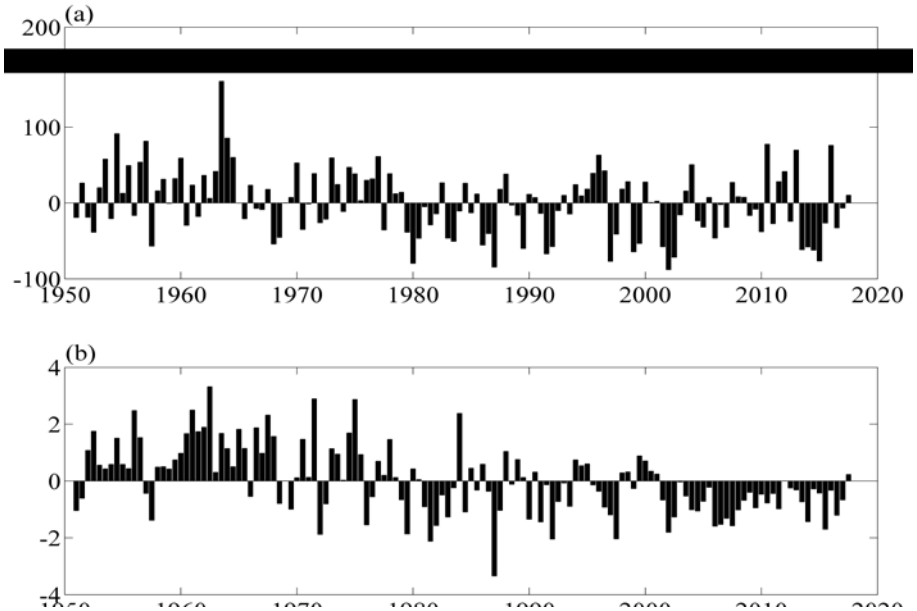

Fig.S2. The anomaly of the precipitation in North China (a) and the anomaly of North African Subtropical High Ridge Position (b) for July-August during 1951-2017.